



A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation

A. I. Requena, L. Mediero, and L. Garrote

Department of Hydraulic and Energy Engineering, Technical University of Madrid, Madrid, Spain

Correspondence to: A. I. Requena (ana.requena@caminos.upm.es)

Received: 20 December 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 15 January 2013

Revised: 8 May 2013 – Accepted: 27 June 2013 – Published: 1 August 2013

Abstract. A multivariate analysis on flood variables is needed to design some hydraulic structures like dams, as the complexity of the routing process in a reservoir requires a representation of the full hydrograph. In this work, a bivariate copula model was used to obtain the bivariate joint distribution of flood peak and volume, in order to know the probability of occurrence of a given inflow hydrograph. However, the risk of dam overtopping is given by the maximum water elevation reached during the routing process, which depends on the hydrograph variables, the reservoir volume and the spillway crest length. Consequently, an additional bivariate return period, the so-called routed return period, was defined in terms of risk of dam overtopping based on this maximum water elevation obtained after routing the inflow hydrographs. The theoretical return periods, which give the probability of occurrence of a hydrograph prior to accounting for the reservoir routing, were compared with the routed return period, as in both cases hydrographs with the same probability will draw a curve in the peak-volume space. The procedure was applied to the case study of the Santillana reservoir in Spain. Different reservoir volumes and spillway lengths were considered to investigate the influence of the dam and reservoir characteristics on the results. The methodology improves the estimation of the Design Flood Hydrograph and can be applied to assess the risk of dam overtopping.

probability of occurrence of the hydrological event (Chebana and Ouarda, 2011). Moreover, the full hydrograph is of interest in the case of dam design, as the inflow peak is transformed into a different outflow peak during the routing process in the reservoir. Therefore, due to the multivariate nature of flood events, a multivariate frequency analysis of random variables such as flood peak, volume and duration is required to design some structures like dams.

National laws and guidelines usually fix a given return period for dam design. Among others, France uses a return period of 1000 to 10 000 yr depending on the dam typology; Austria fixes a return period of 5000 yr and Spain uses a return period of 500 to 10 000 yr depending on the dam typology and its downstream vulnerability (Minor, 1998; Rettemeier and Köngeter, 1998). However, they do not specify whether it is the return period of either the peak, or hydrograph volume or the entire hydrograph. Moreover, the risk related to a specific event can be over- or underestimated if only the univariate return period of either the peak or volume is analysed (Salvadori and De Michele, 2004; De Michele et al., 2005). In addition, the return period should be defined in terms of risk of either dam overtopping or downstream damages, instead of in terms of natural probability of occurrence of floods, to take into account the influence of reservoir and dam characteristics on the flood hydrograph routing process (Mediero et al., 2010). In the case of risk of dam overtopping, the maximum water level reached during the routing process should be used to define the return period. Nevertheless, the relationship between a given inflow hydrograph and its maximum water level is not straightforward, as it depends on the reservoir volume and the spillway crest length. Consequently, the routing process has to be studied in each dam.

1 Introduction

Univariate flood frequency analyses have been carried out widely, focusing on the study of flood peaks, which are used for designing most of hydraulic structures. However, univariate frequency analyses do not procure a full evaluation of the

Hence, a multivariate analysis on flood variables should be conducted to obtain both the natural probability of occurrence of a flood and the return period of a flood in terms of risk of dam overtopping. This analysis has been traditionally undertaken through the use of a stochastic weather generator and continuous rainfall-runoff models (Calver and Lamb, 1995; Cameron et al., 1999; Blazkova and Beven, 2004). Although this approach has proven very successful, it is computationally very demanding, especially if extreme events are the focus of the analysis and an estimation of uncertainty is required. Copula models are a valid alternative, because they allow generating arbitrarily long series to extend the observed hydrological data with less computational effort than continuous rainfall-runoff models.

Traditional multivariate techniques assume that the marginal distributions should come from the same family of distributions and the dependence between variables follows a linear relationship. However, drawbacks arise because these assumptions could not be satisfied by the dependence structure of flood variables. Copula models can avoid these difficulties. A copula is a function that connects univariate distribution functions to a multivariate distribution function describing the dependence among correlated variables (Nelsen, 1999). The main advantage of copulas is that univariate marginal distributions can be defined independently from the joint behaviour of the variables involved. Hence, a copula allows for modelling the dependence structure of random variables regardless the family that the marginal distributions belong to. Besides, joint return periods can be easily estimated from copulas, which represents an additional benefit as the study of joint return periods is essential to flood frequency analysis.

The theory of copulas is based on the Sklar's theorem (Sklar, 1959), which in the case of a bivariate case can be written in the form:

$$H(x, y) = C\{F(x), G(y)\}, x, y \in \mathfrak{R}, \quad (1)$$

where $H(x, y)$ is the joint cumulative distribution function of the random variables X and Y , $F(x)$ and $G(y)$ are the marginal distribution functions of X and Y , respectively, and the mapping function $C: [0, 1]^2 \rightarrow [0, 1]$ is the copula function.

Further details about copulas can be found in Joe (1997), Nelsen (1999) and Salvadori et al. (2007).

Although copula models have been extensively applied in other fields such as finance, they have been only recently applied to model hydrological events such as floods, storms and droughts. Overall, the Archimedean and extreme value copula families are the most used in modelling flood variables. The Archimedean copulas can be constructed easily and, as a great deal of copulas belongs to this family, a broad kind of dependence can be considered. Some authors used Archimedean copulas such as the Frank copula (Favre et al., 2004) or the Clayton copula (Shiau et al., 2006) to characterise the dependence structure between peak and volume

variables. Meanwhile, extreme value copulas have the advantage that they are able to connect the extreme values of the studied variables, which is very important in flood frequency analysis. A lot of authors considered the Gumbel copula as the copula that best represents the relation between peak and volume (Zhang and Singh, 2006, among others).

But selection of the copula model that best fits the observed data is not a trivial issue. Some works have been carried out in recent years regarding the steps required to select a copula model. Using a small sample, Genest and Favre (2007) described different aspects to take into account in the process of studying the dependence between two random variables, in order to identify the appropriate copula model. The importance of considering upper tail dependence in copula selection was emphasised by Poulin et al. (2007), in order not to underestimate the flood risk, as the upper tail dependence is related to the degree of dependence between the extreme values of the variables involved in the study. Thereby, Chowdhary et al. (2011) indicated the steps needed to select the best copula model taking into account the tail dependence in the decision process.

Different bivariate return periods estimated by copulas have been developed in the last years. Salvadori and De Michele (2004) studied the unconditional and conditional return periods of hydrological events using copulas, focussing on the joint return period in which either x or y are exceeded (primary return period) and on the joint return period in which both x and y are exceeded. An additional return period was also introduced, the secondary return period (also called the Kendall return period), which is associated with the realisation of dangerous events for the dam. As it is linked to the primary return period, it can be understood as the mean inter-arrival time of the events with a primary return period over a threshold (called critical events). Authors such as Shiau et al. (2006) also applied joint return periods to study the bivariate flood frequency analysis of peak and volume. A comparison of the different return period approaches for the estimation of design events has been recently shown by Gräler et al. (2013). Other studies have been carried out regarding multivariate flood frequency analysis using copulas (Grimaldi and Serinaldi, 2006; Serinaldi and Grimaldi, 2007; Zhang and Singh, 2007).

Other authors have studied dam safety more in depth using copulas. De Michele et al. (2005) utilised the Gumbel copula to generate peak-volume pairs, in order to verify that the maximum water level reached at the dam by the generated hydrographs was below the crest level. Klein et al. (2010) presented a methodology to classify floods regarding the hydrological risk, estimating the probability of occurrence of peak and volume via a copula model. The maximum water level reached at the dam by each flood was estimated, to graphically analyse the relation between this level and the value of the primary return period for each event.

In this paper, a bivariate flood frequency analysis was carried out by a copula model to conduct a comparison between

the return periods that estimate the natural probability of occurrence of floods and a return period defined in terms of risk of dam overtopping (so-called the routed return period), to assess the influence of the routing process on them. Theoretical return periods based on the joint probability of occurrence were estimated from the fitted copula. In addition, the fitted copula was used to generate a large set of synthetic peak-volume flood pairs in the catchment. Synthetic hydrographs were generated to ascribe a shape to the synthetic peak-volume pairs. The set of synthetic hydrographs was routed through the reservoir to obtain the maximum water level reached at the dam during the routing process, in order to assess the hydrological risk of dam overtopping. The routed return periods based on the risk of dam overtopping were estimated. Furthermore, a sensitivity analysis on the reservoir volume and spillway crest length was carried out, to investigate how the flood variables control the routed return period depending on the dam and reservoir characteristics. The methodology was applied to the Santillana reservoir in Spain.

The structure of the paper is the following: the proposed methodology is shown in Sect. 2. Section 3 presents the case study. Then, the results obtained after applying the procedure are included in Sect. 4. Conclusions are introduced in Sect. 5.

2 Methodology

In this section the proposed methodology is presented. First, the steps followed to select the copula model from observed data are described. Then, the theoretical joint return periods are briefly described. Thirdly, the procedure to generate synthetic hydrographs is presented. Finally, the procedure to obtain the routed return period in terms of risk of dam overtopping is offered.

2.1 Copula selection

Identification of the copula that best fits the observations is required, as several families of copulas exist. The copula that best represents the dependence structure between variables will be the most appropriate. The steps involved in selecting the appropriate copula model are (i) dependence evaluation; (ii) parameter estimation method; (iii) goodness-of-fit tests; and (iv) tail dependence assessment.

2.1.1 Dependence evaluation

A dependence analysis among correlated random variables is conducted to determine if some kind of dependence can be deduced from the data. It can be carried out by graphical analyses or dependence measures. A graphical analysis of dependence can be displayed by the scatter plot of the pairs $(R_i/(n+1), S_i/(n+1))$ derived from the observed data pairs (X_i, Y_i) (where R_i is the rank of X_i among X_1, \dots, X_n and S_i is the rank of Y_i among Y_1, \dots, Y_n , being $i = 1, \dots, n$ and

n the observed record length), and by other two rank-based scatter plots: the Chi-plot and the K-plot.

The Chi-plot displays a measure of location of an observation regarding the whole of the observations (λ_i) against a measure of the well-known Chi-square test statistic for independence (χ_i). Consequently, the larger the distance between the points and the zero value in the y axis, the larger is the dependence. The dependence is positive if the points are above the upper control limit and negative if they are located below the lower control limit (Fisher and Switzer, 1985, 2001).

The K-plot relates the order statistics (H_i) estimated from the observed data to the expected value of these statistics ($W_{i:n}$) generated under the null hypothesis of independence between the marginal distributions. Therefore, the larger the distance between the points and the diagonal line, the larger the dependence. Hence, if the points are located above this line, the dependence is positive. On the other hand, if the points are below this line, the dependence is negative (Genest and Boies, 2003).

Besides, dependence measures are needed to procure a quantitative value of the dependence relation between variables. For this purpose, the Spearman's rho and Kendall's tau rank-based non-parametric measures of dependence are adopted and its associated p values are estimated (independence between variables is rejected when the p value is less than 0.05, further details about the p value estimation can be found in Genest and Favre, 2007). The result of this evaluation provides an idea of the type of copula to be considered in the study, since each copula supports a particular range of dependence parameter. Michiels and De Schepper (2008) provides ranges of admissible Kendall's tau to different copulas for the bivariate case. Therefore, the number of feasible copulas can be reduced using the Kendall's tau value.

2.1.2 Parameter estimation method

The estimation of the parameter (θ) of the copula family $C_\theta(u, v)$ that best fits the data can be performed through different methods. A first group consists of rank-based methods, in which the parameter estimation is independent of the marginal functions, such as the method based on the inversion of a non-parametric dependence measure (e.g. the inversion of Kendall's tau dependence measure) and the maximum pseudo-likelihood method (MPL). The former is related to the method of moments, while the MPL is a modification of the traditional maximum likelihood method, in which the empirical marginal distributions are used instead of the parametric marginal distributions. Other methods certainly depend on the marginal distributions, such as the inference function for margins (IFM) method proposed by Joe and Xu (1996). There is no consensus, but a large number of authors defend the use of the rank-based estimation methods. Supporting this position, Kim et al. (2007) argue that IFM methods are non-robust against misspecification of the marginal distributions, as the parameter estimation depends

on the choice of the univariate marginal distributions and can be affected if such models do not fit adequately. Consequently, in the present work two rank-based methods were used: the inversion of Kendall's tau method and the MPL method.

2.1.3 Goodness-of-fit tests

The aim of a goodness-of-fit test is selecting the copula that best represents the dependence structure of observed variables. Graphical tools and formal tests are provided to achieve this purpose.

A first idea of the behaviour of the copulas can be drawn via a scatter plot, where a synthetic sample of pairs generated from each copula of study (U_{1j}, U_{2j}) , being $j = 1, \dots, m$ and m the sample size, is compared to the pairs related to the observations $(R_i/(n+1), S_i/(n+1))$. Another useful graph can be elaborated by fitting the marginal distributions of the random variables in order to transform the pairs generated from the copula into their original units (X_j, Y_j) , following Eq. (2):

$$(X_j, Y_j) = (F^{-1}(U_{1j}), G^{-1}(U_{2j})), \quad (2)$$

where F^{-1} and G^{-1} are the inverses of the marginal distributions functions F and G , respectively.

A third graph is the generalized K-plot, based on the procedure introduced by Genest and Rivest (1993), in which a comparison of parametric and non-parametric estimates of $K(t)$ is conducted, being $K(t)$ (the so-called Kendall function) the probability that the copula function is equal or smaller than $t \in [0, 1]$, i.e. the cumulative distribution of the copula value. This widely used procedure has been specially designed for Archimedean copulas, so there are circumstances in which a goodness-of-fit test based on it is not consistent. This is the case of extreme value copulas, as $K(t)$ is the same for all of extreme value models (Genest et al., 2006).

Although graphical tools provide a general notion of the goodness-of-fit, formal tests are needed to quantify it. Several procedures have been proposed in the last years. Genest et al. (2009) show a review and analyse various rank-based procedures. These procedures are classified in three groups: tests based on the empirical copula, tests based on Kendall's transform and tests based on Rosenblatt's transformation. The results indicate that overall, the Cramér-von Mises statistic (S_n based on the empirical copula) has the best behaviour for all copula models, allowing to differentiate among extreme value copulas. It also emphasised the importance of calculating the p value associated to the goodness-of-fit test to formally assess whether the selected model is suitable. The p value is obtained through a parametric bootstrap-based procedure that was validated (Genest and Remillard, 2008). The S_n statistic can be written as

$$S_n = \sum_{i=1}^n \left\{ C_n \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) - C_{\theta_n} \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \right\}^2, \quad (3)$$

where

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n 1 \left(\frac{R_i}{n+1} \leq u, \frac{S_i}{n+1} \leq v \right), \quad u, v \in [0, 1]. \quad (4)$$

Here C_n is the empirical copula (a non-parametric rank-based estimator of the unknown copula), C_{θ_n} the parametric copula with the parameter previously estimated from the observed data and $1(A)$ the indicator function of the set A .

The S_n statistic based on the empirical copula was the goodness-of-fit test utilised in the present paper. The statistic value is used to classify the copula models as the p value can only be utilised to accept or reject each copula model (Salvadori and De Michele, 2011). Consequently, the selected copula should have the lower value of the statistic with an admissible p value (i.e. larger than 0.05).

2.1.4 Tail dependence assessment

Once the goodness of fit of the copula to observed data is assessed, a tail dependence assessment is conducted to evaluate the copula behaviour for high return periods. The idea of tail dependence is connected with the degree of dependence in the upper-right-quadrant tail or lower-left-quadrant tail of a bivariate distribution. In this work, more attention is paid on the upper tail dependence (Serinaldi, 2008), not being relevant the analysis of the lower tail dependence due to the focus on the frequency analysis of extreme flood events for dam safety analysis. Upper tail dependence is associated with the capacity to link extreme flood peaks to extreme volumes. This measure can be quantified by the upper tail dependence coefficient, λ_U . Its general expression is given in Eq. (5).

$$\lambda_U = \lim_{w \rightarrow 1^-} P(F(x) > w | G(y) > w) \quad (5)$$

Equation (5) shows that the upper tail dependence is related to the probability of the marginal distribution of x being larger than the threshold $w \in [0, 1]$, when the marginal distribution of y being also larger than that threshold.

One of the main advantages of copulas is that both the upper and lower tail dependences are inherent to the copula model and depend on its parameter. Therefore, the previous measure can also be estimated using copulas by means of Eq. (6).

$$\lambda_U^C = \lim_{w \rightarrow 1^-} \frac{1 - 2w + C_{\theta}(w, w)}{1 - w} \quad (6)$$

As the aim of this section is to identify the copula models that suitably reproduce the dependence in the extremes, the upper tail coefficient obtained from each copula should be compared to the coefficient estimated from the observed data. Initially, a graphical analysis of the tail dependence of

the observed data is done based on the Chi-plot (Abberger, 2005). A non-parametric estimator of the upper tail dependence coefficient is also obtained in order to be compared quantitatively to the upper tail dependence coefficient of each selected copula. In the present study the considered estimator is $\hat{\lambda}_U^{\text{CFG}}$ (Eq. 7). The estimator was proposed by Frahm et al. (2005). Among others, the estimator has been applied by Serinaldi (2008).

$$\hat{\lambda}_U^{\text{CFG}} = 2 - 2 \exp \left[\frac{1}{n} \sum_{i=1}^n \log \left\{ \frac{\sqrt{\log \frac{1}{u_i} \log \frac{1}{v_i}}}{\log \frac{1}{\max(u_i, v_i)^2}} \right\} \right] \quad (7)$$

The estimator is based on the assumption that the empirical copula can be approximated by an extreme value copula. It also works well when this hypothesis is not fulfilled, except in the case that the real upper tail dependence is null. By means of this analysis, copulas that reproduce properly the dependence in the extremes are identified. Because of a good upper tail dependence fit does not mean a good whole data fit, the assessment of the tail dependence is developed at this point and not before. Therefore, the best copula is the copula which represents properly the dependence structure of the variables peak and volume and allows to study adequately the extreme events.

2.2 Joint return periods

Different joint return periods estimated by the fitted copula have been developed for the case of a bivariate flood frequency analysis. The joint return period $T_{X,Y}^{\vee}$, so-called OR return period, (in which the threshold x or y are exceeded by the respective random variables X and Y) and $T_{X,Y}^{\wedge}$, so-called AND return period, (in which the threshold x and y are exceeded by the respective random variables X and Y) were considered. $T_{X,Y}^{\vee}$ is also known as the primary return period. Using copulas these joint return periods are expressed as

$$T_{X,Y}^{\vee} = \frac{\mu_T}{P(X > x \vee Y > y)} = \frac{\mu_T}{1 - C(F(x), G(y))} \quad (8)$$

$$T_{X,Y}^{\wedge} = \frac{\mu_T}{P(X > x \wedge Y > y)} = \frac{\mu_T}{1 - F(x) - G(y) + C(F(x), G(y))} \quad (9)$$

where $C(F(x), G(y)) = P(X \leq x \wedge Y \leq y)$ and μ_T is the mean inter-arrival time between two successive events ($\mu_T = 1$ for maximum annual events). Besides, the following inequality is always fulfilled:

$$T_{X,Y}^{\vee} \leq \min[T_X, T_Y] \leq \max[T_X, T_Y] \leq T_{X,Y}^{\wedge}, \quad (10)$$

where T_X and T_Y are the univariate return periods.

An additional return period is also studied, the secondary return period ρ_t^{\vee} . The secondary return period (Eq. 11) is associated with the primary return period, as it can be defined as the mean inter-arrival time of an event with a primary return period larger than a threshold $\vartheta(t)$. That is to say, it is related to the probability of occurrence of an event in the

area over the copula level curve of value t (Salvadori and De Michele, 2004).

$$\rho_t^{\vee} = \frac{\mu_T}{1 - K(t)} \quad (11)$$

The three joint return periods can be easily obtained using copulas thanks to their formulation. Once the copula selection is completed, the level curves of the fitted copula will be the curves where the events with the same probability of occurrence are located, as the copula value indicates the probability of both x and y are not exceeded.

2.3 Synthetic hydrograph generation

Synthetic hydrographs were estimated from flood peak-volume pairs obtained by means of the selected copula, in order to be routed through the reservoir. A set of observed hydrographs was used as a random sample to ascribe a hydrograph shape to each peak-volume pair. The procedure is the following (Mediero et al., 2010): (i) The ratio between peak and volume is calculated for each peak-volume pair generated by the copula; (ii) The shape of the observed hydrograph with the closest ratio is selected; and (iii) The synthetic peak value is utilised to rescale the selected hydrograph and the synthetic volume is adjusted by modifying the hydrograph duration. A set of 100 000 synthetic hydrographs was generated by this procedure, to have a large sample to study high return periods.

2.4 Routed return period in terms of risk of dam overtopping

The set of generated synthetic hydrographs was routed through the reservoir to assess the risk of dam overtopping by the maximum water level (MWL) reached during the routing process. The analysis is based on the assumption that hydrological risk at the dam is related to MWL, as a return period should be defined in terms of acceptable risk to the structure. Consequently, the routed return period related to the risk of dam overtopping (T_{dam}) can be calculated as the inverse of the probability to exceed a MWL any given year (p_{exc}):

$$T_{\text{dam}} = \frac{1}{p_{\text{exc}}}. \quad (12)$$

For this purpose, the univariate frequency curve of MWL is obtained and the MWL for a given return period is estimated (Fig. 1a). Furthermore, hydrographs with different combinations of peak and volume can lead to a similar MWL, which implies a similar risk of dam overtopping and, consequently, they can be considered to have the same return period. In the bivariate case, these hydrographs can be represented by a curve in the peak-volume space (Mediero et al., 2010). Thereby, return period curves that represent the same risk to the dam are obtained as curves in the peak-volume space (Fig. 1b).

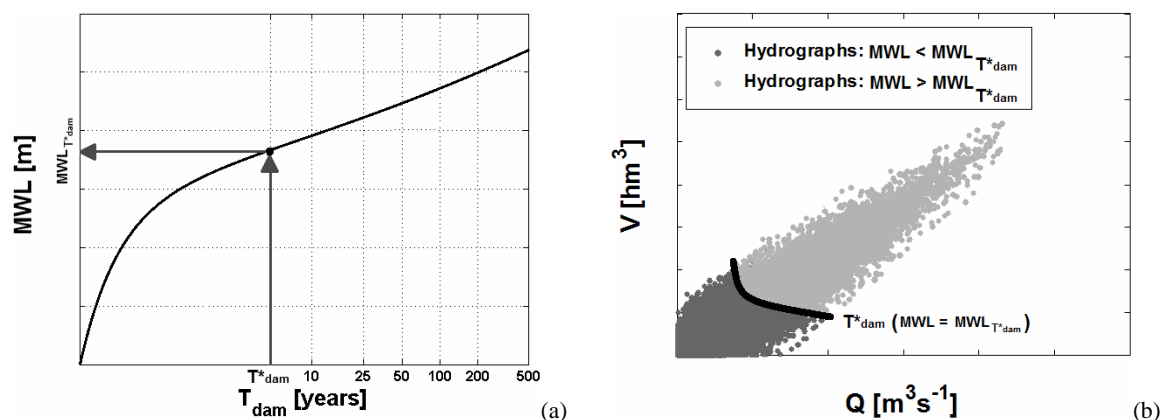


Fig. 1. Procedure to obtain the routed return period curve that represent the risk of dam overtopping (T_{dam}). (a) Estimation of the MWL for the given return period T_{dam}^* ($MWL_{T_{\text{dam}}^*}$) from the frequency curve of MWL. (b) Estimation of the routed return period curve by selecting the hydrographs that reach a given water level equal to $MWL_{T_{\text{dam}}^*}$.

In addition, the influence of reservoir volume and spillway crest length on the shape of these curves was analysed. Different reservoir volumes and spillway crest lengths were considered in order to study the hydrological risk at the dam in different cases.

Finally, the curves that represent the probability of occurrence of floods regarding the theoretical joint return periods based on copulas are compared to the curves that represent the risk of dam overtopping by the routed return period, to assess their similarity and improve the estimation of the Design Flood Hydrograph.

The present study has been carried out by means of both, the commercial software Matlab (Matlab 2009a, The MathWorks, Inc.) and the free software R (R Development Core Team, 2012). Specifically, the use of the R package copula (Kojadinovic and Yan, 2010) is highlighted.

3 Case study

The Santillana reservoir was selected as a case study. It is located in the central west of Spain on the Manzanares River, which belongs to the Tagus basin (Fig. 2). The Santillana reservoir has a drainage area of 325.6 km^2 and a reservoir volume of 92 hm^3 . The elevation of the spillway crest (E) is 889 m , being the flooded area at the spillway crest height of 5.35 km^2 and the reservoir volume up to the spillway crest of 48.9 hm^3 . The dam is an earthfill embankment with a height of 40 m and a crest length (L) of 1355 m . The controlled spillway has a 12 m gate and a maximum capacity of $300\text{ m}^3\text{ s}^{-1}$. A set of 41 yr of observed data was recorded at the reservoir. Observed data are composed of pairs of maximum annual flood peak (Q) and flood volume (V), the latter being the volume of the hydrograph associated to the event with the annual maximum flood peak. As this work is based on the flood frequency curve of maximum annual

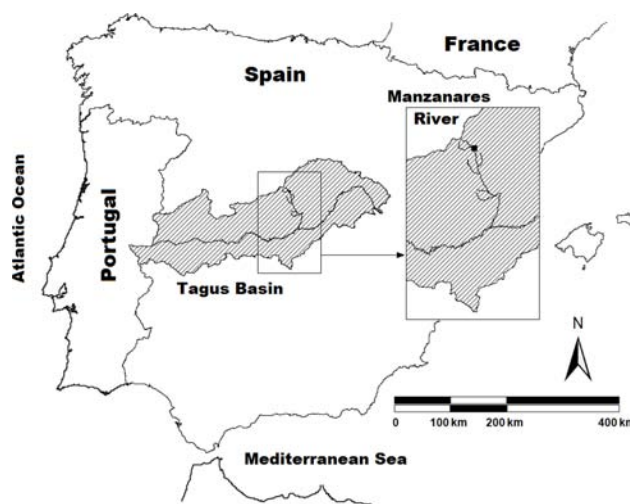


Fig. 2. Location of the Santillana reservoir.

peak discharges, for the sake of consistency, maximum annual flood volumes were assumed to be linked to hydrographs corresponding to the annual maximum peaks.

The marginal distributions for both variables were fitted to a Gumbel distribution, estimating parameters by the L-moments estimation method (Table 1). A prior study carried out in Spain showed that in this region, the Gumbel and the generalised extreme value marginal distributions are appropriate for fitting the annual maximum flood peaks (Jiménez-Álvarez et al., 2012). Consequently, for the sake of simplicity, the Gumbel distribution was considered to fit the data. With the aim of validating this assumption, the Kolmogorov–Smirnov test was applied to the flood peak and volume data set. The p values obtained were larger than 0.05 (0.8426 and 0.9271, respectively), proving that the hypothesis can be accepted.

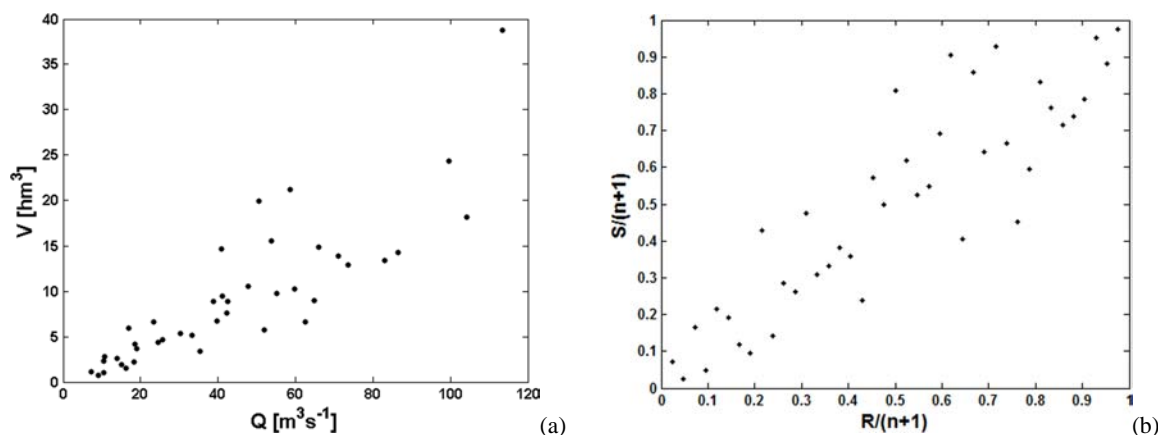


Fig. 3. Scatter plot. (a) Observed data (Q_i , V_i). (b) Ranks ($R_i/(n+1)$, $S_i/(n+1)$) derived from the observed data.

Table 1. Location parameter (μ) and scale parameter (γ) of Gumbel distributions for the variables of peak (Q) and volume (V).

Variable	μ	γ
Q	30.47	22.69
V	5.87	5.70

Table 2. Rank-based non-parametric measures of dependence: Spearman's rho (ρ) and Kendall's tau (τ).

Dependence measure	Value	p value
ρ	0.8899	1.82×10^{-8}
τ	0.7244	2.53×10^{-11}

4 Results

The proposed methodology was applied to the case study.

4.1 Copula selection

Once the univariate marginal distributions are known, the first step consists of studying the dependence between the two random variables: peak and volume. The scatter plot of the observed data is displayed in Fig. 3a. The scatter plot of the pairs ($R_i/(n+1)$, $S_i/(n+1)$) derived from the data set shows a positive relation of dependence between variables (Fig. 3b). This fact is also supported by the Chi-plot (Fig. 4a) and the K-plot (Fig. 4b). In the former, the values are located above the upper limit indicating positive dependence. In the latter, the values are plotted over the diagonal line, so positive interaction is also drawn.

The value of the Spearman's rho (ρ) and Kendall's tau (τ) rank-based non-parametric measures of dependence corroborate the results provided by the graphical information. The value of each dependence measure as well as its linked p value are summarised in Table 2.

The set of copulas considered is classified into three classes: Archimedean copulas, extreme value copulas and other families. Ali–Mikhail–Haq, Clayton, Frank and Gumbel copulas belong to the first class, while Galambos, Hüsler–Reiss, and Tawn copula are part of the extreme value copulas family. The Gumbel copula also belongs to the second group. Farlie–Gumbel–Morgenstern and Plackett are included into the last class. The set of feasible copulas was reduced after

testing the admissible range of dependence supported by each one using the Kendall's tau value. As result, Ali–Mikhail–Haq ($\tau \in [-0.1817, 1/3]$), Tawn ($\tau \in [0, 0.4184]$) and Farlie–Gumbel–Morgenstern copula ($\tau \in [-2/9, 2/9]$) were eliminated.

Copula functions and parameter space of the copulas selected in the study are presented in Table 3. The parameter of the copulas is estimated using both rank-based methods, the inversion of Kendall's tau and the MPL method. The standard error (SE) is also obtained for each estimated parameter (Kojadinovic and Yan, 2010). Although SE is not a goodness-of-fit criterion, small values are always desirable. The results in Table 4 show that except in the case of the Clayton and the Plackett copula, the lowest standard error is associated with the inversion of Kendall's tau method. Besides, the standard error linked with the parameter of the extreme value copulas is the smallest.

In all, 100 000 synthetic pairs are generated from each copula. The scatter plot of the synthetic pairs transformed back into its original units using univariate marginal distributions and the observed data are shown in Fig. 5. Only copulas whose parameter is obtained by inversion of Kendall's tau method are drawn. The figure shows that extreme value copulas (Gumbel, Galambos and Hüsler–Reiss) are sharper in the upper right corner while the other copula models are more scattered in this area. This is so because extreme value copulas present positive dependence in the upper tail. The positive lower tail dependence of the Clayton copula can also

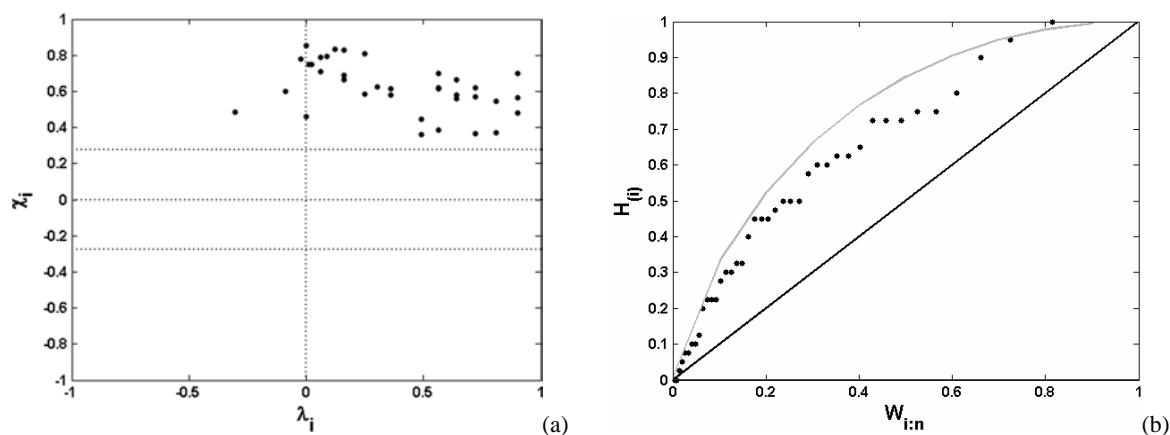


Fig. 4. (a) Chi-Plot and (b) K-plot.

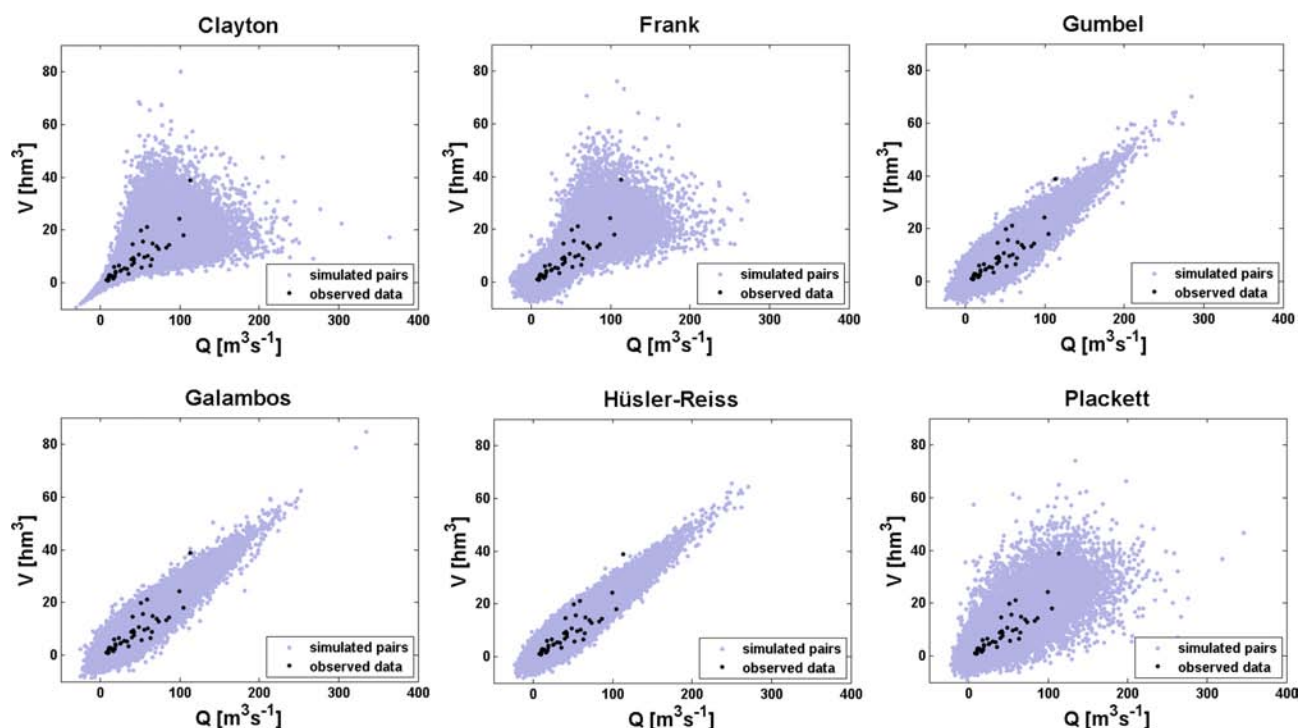


Fig. 5. Scatter plot of 100 000 values generated from the copulas fitted by the inversion of Kendall's tau method and the observed data.

be observed in the graph. Extreme value copulas reproduce the behaviour of the data leaving the largest observation on the edge of the simulated sample, while Clayton, Frank and Plackett copulas include this observation in the set of the generated sample, as their dependence structure in the upper tail is more spread. A further analysis is needed to select the copula that best fits the data.

As expected, the generalized K-plot provides the same information for all of the extreme value copulas (Fig. 6). The distance between parametric (K_{θ_n}) and non-parametric estimate (K_n) of K is greater for extreme value copulas than for the other copula models. Consequently, this analysis shows

that extreme value copulas are slightly worse in terms of fitting to the observed data.

In addition, the S_n goodness-of-fit test based on the empirical copula and its associated p value based on $N = 10\,000$ parametric bootstrap samples (which are also included in Table 4) are estimated for each copula to select the suitable copulas in a formal way. This test shows a good behaviour for all copula families and makes a distinction among extreme value copulas. The parameter of the studied copulas has been estimated using two different methods, the inversion of Kendall's tau and the MPL method. Consequently, not only the comparison among all the S_n values is done, but

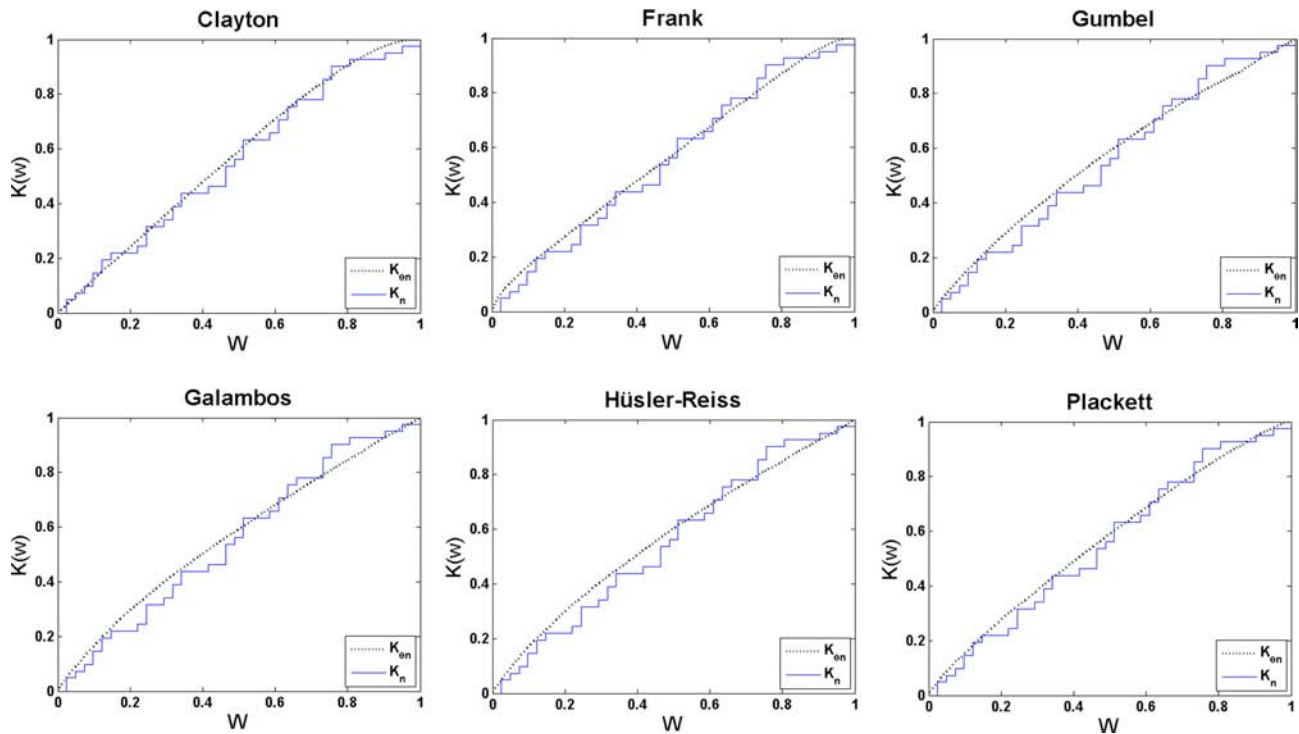


Fig. 6. Comparison between parametric (K_{θ_n}) and non-parametric (K_n) estimate of K , considering the copulas fitted by the inversion of Kendall's tau method.

also the comparison among the S_n values provided by each method, as the estimations provided by different methods can lead to significant differences in results. It can be seen that S_n leads to better results by the inversion of Kendall's tau method than by the MPL method for all copula models. Hence, in this case, the inversion of Kendall's tau method behaves better than the MPL method.

It should be highlighted that although Frank copula with the parameter estimated by the inversion of Kendall's tau method is the most appropriate in terms of fitting the bulk of the observed data (as it has the lower value of S_n and a suitable p value), neither of the remaining copulas could be rejected considering the p value. Therefore, a second analysis is carried out to study the behaviour of the upper part of the distribution by comparing the upper tail dependence of the observed data with the upper tail dependence provided by each copula model. This analysis is used to test which copulas better represent the behaviour of the extreme values of the observed data.

The graphical analysis of the upper tail dependence of the observed data is carried out based on the Chi-plot, only considering the observations located in the upper right corner of the scatter plot (Fig. 7). The analysis indicates that upper tail dependence exists in the data set (what was expected for extreme value data), as the points located in the right edge tend to be far from the zero value of the y axis, which is the independence hypothesis. In addition, Table 5 shows the results

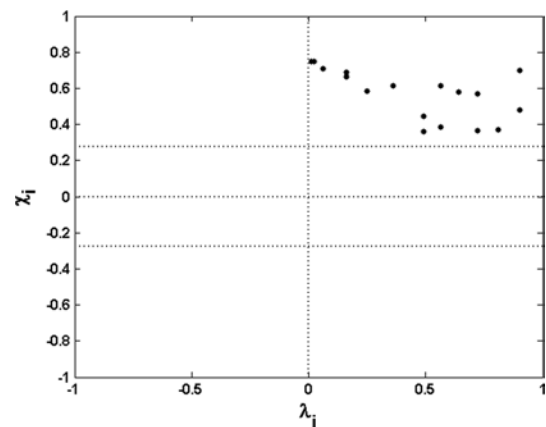


Fig. 7. Upper tail dependence analysis based on Chi-plot.

of the λ_U^C of the studied copulas. The coefficients were estimated using the copula parameter obtained by the inversion of Kendall's tau method, as this method obtained better results. As Fig. 5 announced, only the extreme value copulas show upper tail dependence. The remaining copulas show a null result, as they have by definition an upper tail dependence of zero. Then, the non-parametric estimator of the upper tail dependence coefficient of the observed data obtained by means of Eq. (7), $\hat{\lambda}_U^{CFG} = 0.749$, is compared with the upper tail dependence coefficient of each considered copula. As

Table 3. Copula functions and parameter space of the considered copulas.

Copula	$C_\theta(u, v)$	θ space
Clayton	$\left[\max \left(u^{-\theta} + v^{-\theta} - 1 \right) \right]^{-1/\theta}$	$[-1, \infty) \setminus \{0\}$
Frank	$\frac{-1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right]$	$[-\infty, \infty) \setminus \{0\}$
Gumbel	$\exp \left[- \left(\tilde{u}^\theta + \tilde{v}^\theta \right)^{1/\theta} \right]$	$[1, \infty)$
Galambos	$u v \exp \left[\left(\tilde{u}^{-\theta} + \tilde{v}^{-\theta} \right)^{-1/\theta} \right]$	$[0, \infty)$
Hüsler–Reiss	$\exp \left[-\tilde{u} \Phi \left\{ \frac{1}{\theta} + \frac{\theta}{2} \ln \left(\frac{\tilde{u}}{\tilde{v}} \right) \right\} - \tilde{v} \Phi \left\{ \frac{1}{\theta} + \frac{\theta}{2} \ln \left(\frac{\tilde{v}}{\tilde{u}} \right) \right\} \right]$	$[0, \infty)$
Plackett	$\frac{1}{2} \frac{1}{\theta - 1} \left\{ 1 + (\theta - 1)(u + v) - \left[(1 + (\theta - 1)(u + v))^2 - 4\theta(\theta - 1)uv \right]^{1/2} \right\}$	$[0, \infty)$

Note: $\tilde{u} = -\ln(u)$; $\tilde{v} = -\ln(v)$; and Φ is the univariate standard Normal distribution.

Table 4. Estimated value of the copula parameter (θ_n), copula parameter standard error (SE), Cramér–von Mises goodness-of-fit test (S_n) and p value calculated based on $N = 10\,000$ parametric bootstrap samples, according to the parameter estimation method.

Copula	Parameter estimation method	θ_n	SE	S_n	p value
Clayton	inversion Kendall's tau	5.257	1.202	0.0223	0.3828
	MPL	3.337	1.081	0.0524	0.0565
Frank	inversion Kendall's tau	12.622	2.459	0.0174	0.8402
	MPL	11.774	2.858	0.0202	0.7332
Gumbel	inversion Kendall's tau	3.628	0.601	0.0218	0.3967
	MPL	3.068	0.714	0.0351	0.0649
Galambos	inversion Kendall's tau	2.919	0.602	0.0219	0.3910
	MPL	2.345	0.697	0.0357	0.0603
Hüsler–Reiss	inversion Kendall's tau	3.677	0.684	0.0221	0.3663
	MPL	2.970	0.777	0.0379	0.0568
Plackett	inversion Kendall's tau	54.230	21.699	0.0181	0.7893
	MPL	33.570	17.531	0.0308	0.2967

Table 5. Upper tail dependence coefficient of the considered copulas.

Copula	$\lambda_{U^C}^C(\theta)$	θ_n	$\hat{\lambda}_{U^C}^C$
Clayton	0	5.257	0
Frank	0	12.622	0
Gumbel	$2 - 2^{1/\theta}$	3.628	0.789
Galambos	$2^{-1/\theta}$	2.919	0.789
Hüsler–Reiss	$2 - 2 \Phi \left\{ \frac{1}{\theta} \right\}$	3.677	0.786
Plackett	0	54.230	0

Note: Φ is the univariate standard Normal distribution.

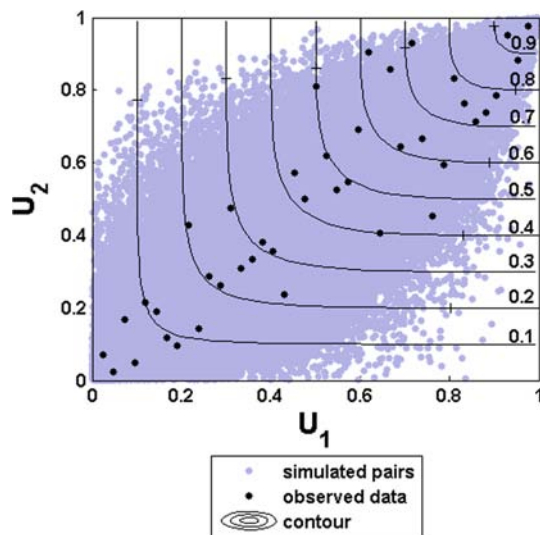
the estimator value is similar to the three values obtained for the extreme value copulas, it can be considered that Gumbel, Galambos and Hüsler–Reiss copulas reproduce suitably the dependence in the upper extreme.

In summary, the best copula should represent properly both, the dependence structure of the observed pairs of peak and volume and the behaviour in the upper part of the distribution. Considering the whole tests, the Gumbel copula was selected as the best copula model. It is an extreme value copula, consequently it takes into account the upper tail dependence and, at the same time, shows a suitable p value representing properly the dependence structure between both variables. Besides, as the Gumbel copula is also an Archimedean copula, it preserves the useful properties of this family, such as the existence of an analytical expression of the Kendall function.

A comparison between a sample generated from the fitted Gumbel copula and the observed data is displayed in Fig. 8. Also, contours of the fitted copula that represent the events with the same probability of occurrence are shown.

Table 6. Comparison between joint return periods associated to the theoretical events with peaks equal to q_T and volumes equal to v_T for $T = 10, 100$ and 1000 yr.

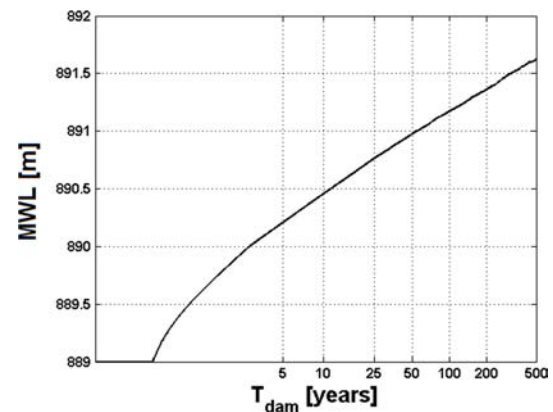
Copula	$T = T_X = T_Y$	q_T ($\text{m}^3 \text{s}^{-1}$)	v_T (h m^3)	t	$T_{X,Y}^\vee$	$T_{X,Y}^\wedge$	$K_{\theta_n}(t)$	ρ_t^\vee
Gumbel	10	81.52	18.69	0.8803	8	12	0.9112	11
	100	134.80	32.08	0.9879	83	127	0.9912	114
	1000	187.12	45.21	0.9988	826	1266	0.9991	1140
Frank	10	81.52	18.69	0.8572	7	17	0.9233	13
	100	134.80	32.08	0.9811	53	891	0.9979	481
	1000	187.12	45.21	0.9980	503	80 226	0.9999	40 448

**Fig. 8.** Comparison between a sample generated from the Gumbel copula and the observed data with the copula contours.

4.2 Joint return periods

Firstly, a brief analysis is conducted to check the results of the copula selection by comparing the risk assumed depending on the selection of a copula model without upper tail dependence (Frank copula) and a copula model with upper tail dependence (Gumbel copula). The joint return periods $T_{X,Y}^\vee$ (Eq. 8), $T_{X,Y}^\wedge$ (Eq. 9) and ρ_t^\vee (Eq. 11) associated to the theoretical events with peaks equal to q_T and volumes equal to v_T for return periods (T) equal to 10, 100 and 1000 yr are estimated for both Gumbel and Frank copula, being q_T and v_T the quantiles obtained from the Gumbel marginal distributions.

The results presented in Table 6 indicate that although $T_{X,Y}^\vee$ linked to the Gumbel copula are higher for all the return periods, $T_{X,Y}^\wedge$ and the ρ_t^\vee are much smaller. It can also be seen that the higher the return period, the larger the differences between joint return periods related to each copula. Therefore, as expected for not being an extreme value copula, the Frank copula underestimates the risk associated to

**Fig. 9.** Frequency curve of MWL for the spillway real set-up ($E = 889$ m, $L = 12$ m).

the joint return periods $T_{X,Y}^\wedge$ and ρ_t^\vee . Hence, this analysis supports the fact that not taking into consideration the upper tail dependence in joint extreme events modelling can lead to an underestimation of the risk (Poulin et al., 2007).

Therefore, once the Gumbel copula was selected as the best copula model, the joint return periods $T_{X,Y}^\vee$, $T_{X,Y}^\wedge$ and ρ_t^\vee were calculated through it.

4.3 Routed return period in terms of risk of dam overtopping

In all, 100 000 annual synthetic hydrographs were estimated by means of the 100 000 peak–volume pairs generated from the Gumbel copula using the procedure explained in Sect. 2.3. The set of hydrographs was routed through the reservoir, which was assumed to be uncontrolled for the sake of simplicity. The frequency curve of the MWL reached was obtained for the spillway real set-up (an elevation of the spillway crest of 889 m and a spillway length of 12 m) (Fig. 9). Maximum water level quantiles for a given return period were estimated easily from this frequency curve (Table 7). Return period curves in the peak–volume space regarding the risk to the dam were obtained as the hydrographs that lead to a similar MWL. Thereby, Fig. 10 shows the

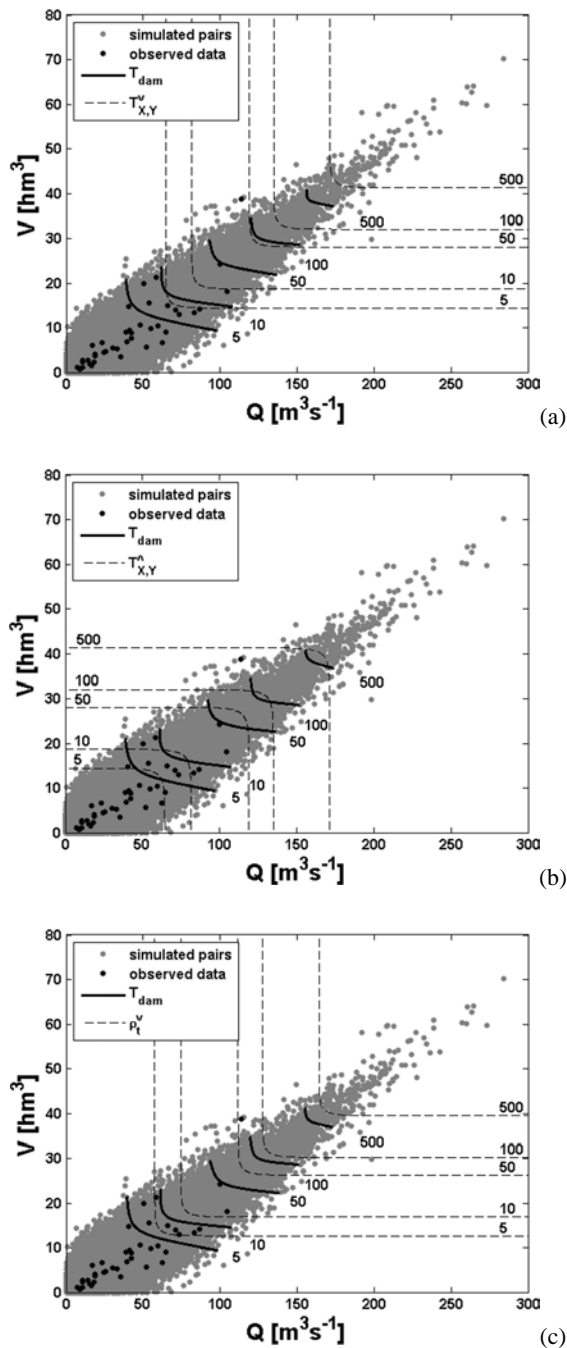


Fig. 10. Comparison among return periods curves that represent the risk to the dam T_{dam} and joint return periods curves (a) $T_{X,Y}^V$, (b) $T_{X,Y}^A$ and (c) ρ_t^V , for the spillway real set-up ($E = 889$ m, $L = 12$ m).

comparison among the different curves that represent different risks: (i) the theoretical curves associated with the joint return periods $T_{X,Y}^V$, $T_{X,Y}^A$ and ρ_t^V estimated from the fitted copula, which are probabilistic based and show a supposed risk of a hydrograph being harmful prior to accounting for the effect of the routing reservoir process on it, and (ii) the curves

Table 7. Maximum water level reached for different return periods (MWL) associated to the probability of exceeding a water level for $E = 889$ m and $L = 12$ m.

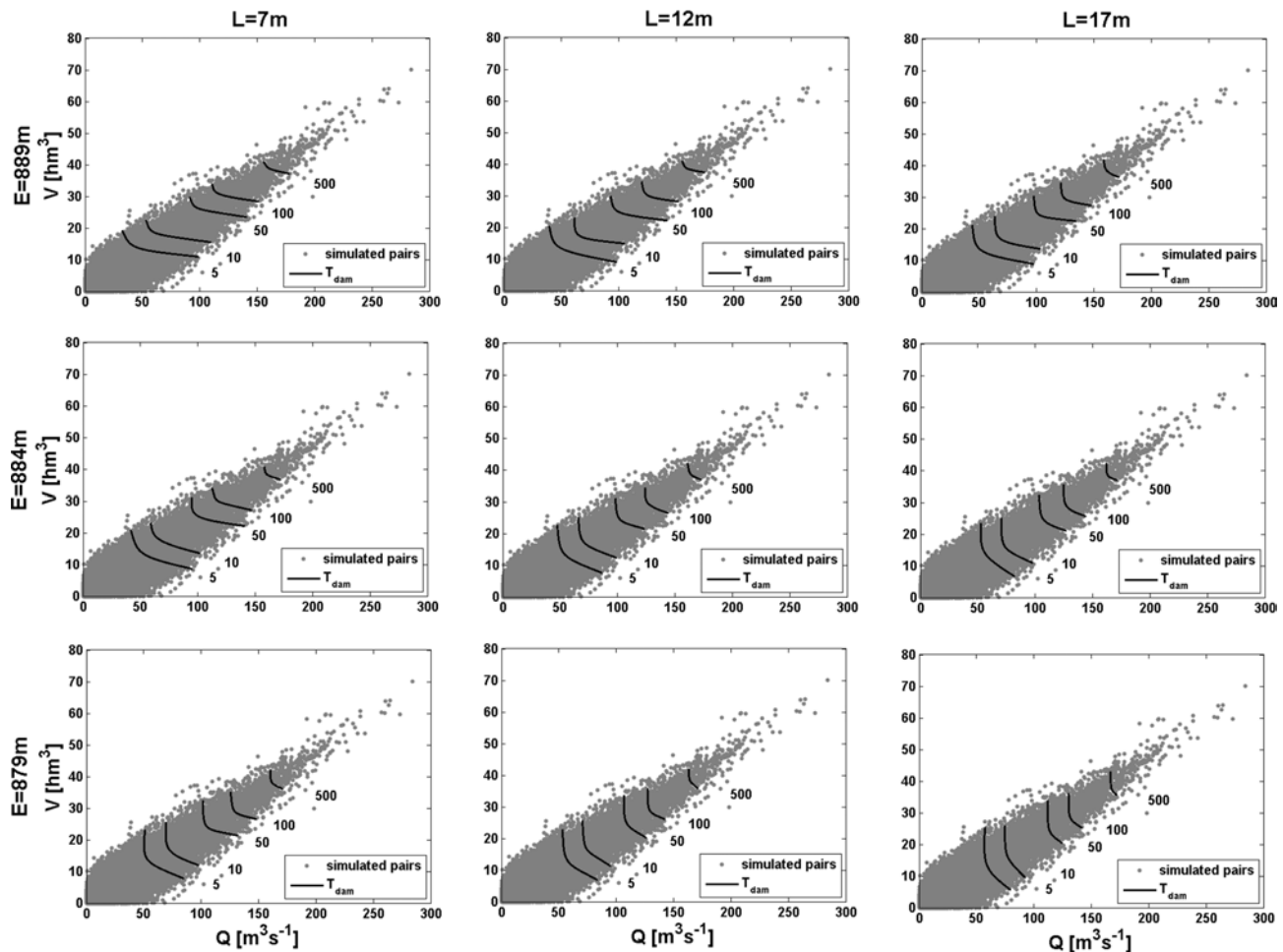
T_{dam} (yr)	p_{exc}	MWL (m)
5	0.2	890.21
10	0.1	890.46
50	0.02	890.98
100	0.01	891.17
500	0.002	891.62

that represent the risk of dam overtopping by the routed return period (T_{dam}), which are related to a real risk for the structure, as they are obtained based on the MWL reached. Figure 10 provides useful information about observed and predicted events. It can be seen that the secondary return period curves are the most similar to the return period curves that represent the risk to the dam. The secondary return period is linked to the probability that an event with a copula value higher than t occurs. The routed return period is calculated from the probability of exceeding a water level. As an example, Table 8 summarises this information for two specific events with a copula value of 0.9 and 0.99. The results fulfil Eq. (10).

Once the different curves were compared, a further analysis is carried out on the return period related to the risk to the dam, in order to assess its sensitivity. Figure 11 displays the return period curves related to the risk of dam overtopping for different reservoir volumes given by reservoir elevations of 879, 884 and 889 m and spillway lengths of 7, 12 and 17 m. It can be appreciated that the higher the reservoir volume, the more horizontal the curves, while the longer the spillway length, the steeper the curves. Thereby, the most horizontal curve is associated with the highest reservoir volume ($E = 889$ m) and the shortest spillway length ($L = 7$ m), while the steepest curve is linked to the smallest reservoir volume ($E = 879$ m) and the longest spillway length ($L = 17$ m). This is caused by flood control properties in a reservoir: the higher the reservoir volume, the greater the capacity to store hydrograph water volume temporarily and, consequently, the higher the attenuation of the flood peak. In this case, the hydrographs that have more influence on the risk to the dam, or the most dangerous hydrographs, are characterised by a high volume. Consequently, T_{dam} is mostly given by the marginal return period of hydrograph volumes (the curves are more horizontal). On the other hand, the smaller the reservoir volume, the lower the capacity to store water temporarily and the smaller the attenuation of the flood peak. In this case, the hydrographs that have more influence on the risk to the dam are characterised by a high flood peak and T_{dam} is mostly given by the marginal return period of flood peaks (the curves are steeper).

Table 8. Example of comparison among return periods. Two simulated events with a copula value of 0.9 and 0.99 are considered. All return periods are expressed in years.

Q ($\text{m}^3 \text{s}^{-1}$)	V (h m^3)	t	$T_{X,Y}^\vee$	T_X	T_Y	$T_{X,Y}^\wedge$	$K_{\theta_n}(t)$	ρ_t^\vee	MWL (m)	T_{dam}
90.52	19.12	0.9	10	15	11	16	0.9261	14	890.53	13
136.41	34.43	0.99	100	107	151	168	0.9927	138	891.34	180

**Fig. 11.** Comparison among return period curves related to the risk to the dam depending on the reservoir volume and the spillway length.

In the case of the spillway length, the shorter the spillway length, the lower the capacity to discharge and the higher the need of a larger capacity to store water temporarily. The hydrographs that lead to higher maximum water levels will have greater volumes and T_{dam} is mostly given by the marginal return period of hydrograph volumes.

As available flood control volume and spillway crest length vary, so do the location and range of critical values per margin. For further analysis of the results, a detailed comparison is shown in Fig. 12. The comparison is focused on the curves related to the risk of dam overtopping when T_{dam} equals 5 and 50 yr. Figure 12a shows the sensitivity

to reservoir volume for the central value of spillway crest length (12 m), while Fig. 12b shows the sensitivity to spillway crest length for the central value of reservoir volume (884 m). The former shows that as the reservoir volume increases, the range of the critical peaks also increases, while the range of the critical volumes decreases. It can also be noted that for higher return periods, the curve slope changes significantly with the reservoir volume, showing a severe reduction of the range of critical volumes, while the range of critical peaks is increased drastically. On the other hand, the latter exhibits that as the spillway crest length increases, the

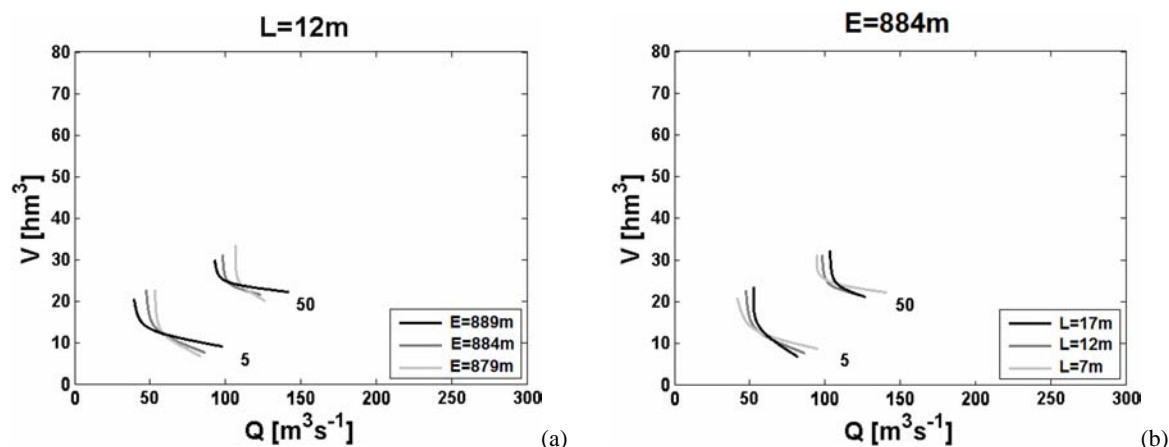


Fig. 12. Detail of the comparison among the routed return period curves when T_{dam} equals 5 and 50 yr, varying with: (a) the reservoir volume and (b) the spillway crest length.

range of the critical peaks decreases, while the range of the critical volume increases.

5 Conclusions

In the present paper a Monte Carlo procedure to obtain the return period linked to the risk of dam overtopping (so-called the routed return period) was carried out by a copula model, comparing it to the probability of occurrence of a flood. For that purpose a bivariate flood frequency analysis of flood peak and volume via a copula model was conducted. The Gumbel copula was found to be the best copula after taking into account the upper tail dependence of the data set. Different joint return period curves were estimated by the copula model fitted to observed data. In addition, a set of synthetic flood hydrographs was generated from the fitted Gumbel copula and was routed through the Santillana reservoir to obtain the maximum water level reached during the routing, as the water level was used as a surrogate of the hydrological risk of dam overtopping. Curves that represent the risk to the dam were obtained as the probability of exceeding a water level. Comparison between both curves for different return periods was carried out. Finally, a sensitivity analysis of the routed return period curves related to the risk to the dam was conducted considering different reservoir volumes and spillway crest lengths.

The results show that tail dependence should be considered in the copula selection to avoid hydrological risk underestimation. The secondary return period curves turned out to be the most similar to the routed return period curves that represent the risk of dam overtopping. These results support the use of the secondary return period in this study. Besides, the use of the secondary return period in preliminary assessment for dam design could be considered adequate, as this return period is independent of reservoir and spillway configuration. The results also hold that the general idea of routing

a large sample through the reservoir has benefits from a conceptual point of view. Accordingly, for detailed design, it could be useful the use of the routed return period obtained by routing a large sample of hydrographs through the reservoir, as it provides a more accurate assessment of the risk of dam overtopping because this return period is specific for the structure under analysis. In summary, in addition to consider the secondary return period, flood hydrographs should also be routed to improve the estimation of the risk of dam overtopping, as there are differences among return periods. It was appreciated that as the available flood control volume increases, the routed return period curves are more dependent on volume (the slope of the return period curves becomes more horizontal). On the other hand, as the spillway length increases, the routed return period curves are more dependent on flood peak (the slope of the return period curves becomes steeper). The sensitivity analysis showed the role played by flood control volume and spillway crest length on the location and ranges per margin of return period curves. The previous results suggest that bivariate analysis could be asymptotically approximated by univariate analysis, for extreme reservoir and spillway characteristics: for a very large flood control volume the dominant variable is hydrograph volume and for a very large spillway length the dominant variable is hydrograph peak. It is also emphasised that the shape of the curves depend not only on the reservoir volume and spillway length, but also on the hydrograph magnitude, given by soil properties, rainfall and physiographic characteristics of the basin.

In conclusion, comparison between bivariate return period curves that represent the risk of dam overtopping (the real risk to the structure) and bivariate joint return period curves that represent the probability of occurrence of a flood event (the theoretical supposed risk without taking into account the structure) provides valuable information about flood control processes in the reservoir. The proposed routed return period

can be useful in dam design, as it improves the estimation provided by the flood event-based return periods, taking into account the dam characteristics that exert a significant influence on the risk of dam overtopping. In addition, this study could be replicated in terms of risk of downstream damages. Therefore, the proposed methodology can procure useful information to estimate the Design Flood Hydrograph.

Acknowledgements. This work has been supported by the Carlos González Cruz Foundation and the project MODEX-Physically based modelling of extreme hydrologic response under a probabilistic approach. Application to Dam Safety Analysis (CGL2011-22868), funded by the Spanish Ministry of Science and Innovation (now the Ministry of Economy and Competitiveness). The authors are also grateful for the financial contribution made by the COST Office grant ES0901 “European procedures for flood frequency estimation (FloodFreq)”.

Edited by: E. Gargouri-Ellouze

References

- Abberger, K.: A simple graphical method to explore tail-dependence in stock-return pairs, *Appl. Financ. Econom.*, 15, 43–51, 2005.
- Blazkova, S. and Beven, K.: Flood frequency estimation by continuous simulation of subcatchment rainfalls and discharges with the aim of improving dam safety assessment in a large basin in the Czech Republic, *J. Hydrol.*, 292, 153–172, 2004.
- Calver, A. and Lamb, R.: Flood Frequency Estimation Using Continuous Rainfall-Runoff Modelling, *Phys. Chem. Earth*, 20, 479–483, 1995.
- Cameron, D. S., Beven, K. J., Tawn, J., Blazkova, S., and Naden, P.: Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty), *J. Hydrol.*, 219, 169–187, 1999.
- Chebana, F. and Ouarda, T.: Multivariate quantiles in hydrological frequency analysis, *Environmetrics*, 22, 63–78, 2011.
- Chowdhary, H., Escobar, L. A., and Singh, V. P.: Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data, *Hydrol. Res.*, 42, 193–216, 2011.
- De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., and Rosso, R.: Bivariate statistical approach to check adequacy of dam spillway, *J. Hydrol. Eng.-ASCE*, 10, 50–57, 2005.
- Favre, A.-C., El Adlouni, S., Perreault, L., Thiémond, N., and Bobée, B.: Multivariate hydrological frequency analysis using copulas, *Water Resour. Res.*, 40, 1–12, 2004.
- Fisher, N. I. and Switzer, P.: Chi-Plots for Assessing Dependence, *Biometrika*, 72, 253–265, 1985.
- Fisher, N. I. and Switzer, P.: Graphical assessment of dependence: Is a picture worth 100 tests?, *Am. Stat.*, 55, 233–239, 2001.
- Frahm, G., Junker, M., and Schmidt, R.: Estimating the tail-dependence coefficient: Properties and pitfalls, *Insur. Math. Econ.*, 37, 80–100, 2005.
- Genest, C. and Boies, J. C.: Detecting dependence with Kendall plots, *Am. Stat.*, 57, 275–284, 2003.
- Genest, C. and Favre, A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask, *J. Hydrol. Eng.-ASCE*, 12, 347–368, 2007.
- Genest, C. and Rémillard, B.: Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, *Ann. I. H. Poincaré-Pr.*, 44, 1096–1127, 2008.
- Genest, C. and Rivest, L.: Statistical Inference Procedures for Bivariate Archimedean Copulas, *J. Am. Stat. Assoc.*, 88, 1034–1043, 1993.
- Genest, C., Quessy, J. F., and Rémillard, B.: Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation, *Scand. J. Stat.*, 33, 337–366, 2006.
- Genest, C., Rémillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study, *Insur. Math. Econ.*, 44, 199–213, 2009.
- Gräler, B., van den Berg, M. J., Vandenbergh, S., Petroselli, A., Grimaldi, S., De Baets, B., and Verhoest, N. E. C.: Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation, *Hydrol. Earth Syst. Sci.*, 17, 1281–1296, doi:10.5194/hess-17-1281-2013, 2013.
- Grimaldi, S. and Serinaldi, F.: Asymmetric copula in multivariate flood frequency analysis, *Adv. Water Resour.*, 29, 1155–1167, 2006.
- Jiménez-Álvarez, A., García-Montañés, C., Mediero, L., Incio, L., and Garrote, J.: Map of maximum flows of intercommunity basins, *Revista de Obras Públicas*, 3533, 7–32, 2012.
- Joe, H.: Multivariate model and dependence concepts, Chapman and Hall, London, 1997.
- Joe, H. and Xu, J. J.: The estimation method of Inference functions for margins for multivariate models, Technical Rep. 166, Department of Statistics, University of British Columbia, British Columbia, 1996.
- Kim, G., Silvapulle, M. J., and Silvapulle, P.: Comparison of semi-parametric and parametric methods for estimating copulas, *Comput. Stat. Data An.*, 51, 2836–2850, 2007.
- Klein, B., Pahlow, M., Hündel, Y., and Schumann, A.: Probability analysis of hydrological loads for the design of flood control systems using copulas, *J. Hydrol. Eng.-ASCE*, 15, 360–369, 2010.
- Kojadinovic, I. and Yan, J.: Modeling multivariate distributions with continuous margins using the copula R package, <http://www.jstatsoft.org/v34/i09/>, *J. Stat. Softw.*, 34, 1–20, 2010.
- Matlab 2009a: The MathWorks, Inc., Natick, Massachusetts, USA, 2009.
- Mediero, L., Jiménez-Álvarez, A., and Garrote, L.: Design flood hydrographs from the relationship between flood peak and volume, *Hydrol. Earth Syst. Sci.*, 14, 2495–2505, doi:10.5194/hess-14-2495-2010, 2010.
- Michiels, F. and De Schepper, A.: A copula test space model how to avoid the wrong copula choice, *Kybernetika*, 44, 864–878, 2008.
- Minor, H. E.: Report of the European R&D Working Group “Floods”, in: Proceedings of the International symposium on new trends and guidelines on dam safety, 17–19 June 1998, Barcelona, Spain, 1541–1550, 1998.
- Nelsen, R. B.: An introduction to copulas, Springer, New York, 1999.

- Poulin, A., Huard, D., Favre, A.-C., and Pugin, S.: Importance of tail dependence in bivariate frequency analysis, *J. Hydrol. Eng.-ASCE*, 12, 394–403, 2007.
- R Development Core Team: R: A language and Environment for Statistical Computing, R Foundation for Statistical Computing, <http://www.R-project.org/>, Vienna, Austria, 2012.
- Rettemeier, K. and Köngeter, J.: Dam safety management: overview of the state of the art in Germany compared to other European countries, in: *Proceedings of the International symposium on new trends and guidelines on dam safety*, 17–19 June 1998, Barcelona, Spain, 55–62, 1998.
- Salvadori, G. and De Michele, C.: Frequency analysis via copulas: Theoretical aspects and applications to hydrological events, *Water Resour. Res.*, 40, 1–17, 2004.
- Salvadori, G. and De Michele, C.: Estimating strategies for multiparameter Multivariate Extreme Value copulas, *Hydrol. Earth Syst. Sci.*, 15, 141–150, doi:10.5194/hess-15-141-2011, 2011.
- Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R.: *Extremes in nature, An approach using copulas*, Springer, Dordrecht, the Netherlands, 2007.
- Serinaldi, F.: Analysis of inter-gauge dependence by Kendall's τ_K , upper tail dependence coefficient, and 2-copulas with application to rainfall fields, *Stoch. Environ. Res. Risk. A*, 22, 671–688, 2008.
- Serinaldi, F. and Grimaldi, S.: Fully nested 3-copula: procedure and application on hydrological data, *J. Hydrol. Eng.-ASCE*, 12, 420–430, 2007.
- Shiau, J., Wang, H., and Tsai, C.: Bivariate frequency analysis of floods using copulas, *J. Am. Water Resour. As.*, 42, 1549–1564, 2006.
- Sklar, A.: Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Stat. Univ. Paris*, 8, 229–231, 1959.
- Zhang, L. and Singh, V. P.: Bivariate flood frequency analysis using the copula method, *J. Hydrol. Eng.-ASCE*, 11, 150–164, 2006.
- Zhang, L. and Singh, V. P.: Trivariate flood frequency analysis using the Gumbel–Hougaard copula, *J. Hydrol. Eng.-ASCE*, 12, 431–439, 2007.